Structure, Vol. 11, 445–457, April, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00046-7

X-Ray and Biochemical Anatomy of an Archaeal XPF/Rad1/Mus81 Family Nuclease: Similarity between Its Endonuclease Domain and Restriction Enzymes

Tatsuya Nishino,[1] Kayoko Komori,[2] Yoshizumi Ishino,[2,3] and Kosuke Morikawa[1,*]

[1]Department of Structural Biology

[2]Department of Molecular Biology

Biomolecular Engineering Research Institute (BERI)

6-2-3 Furuedai

Suita, Osaka 565-0874

Japan

Summary

The XPF/Rad1/Mus81-dependent nuclease family specifically cleaves branched structures generated during DNA repair, replication, and recombination, and is essential for maintaining genome stability. Here, we report the domain organization of an archaeal homolog (Hef) of this family and the X-ray crystal structure of the middle domain, with the nuclease activity. The nuclease domain architecture exhibits remarkable similarity to those of restriction endonucleases, including the correspondence of the GDX_nERKX_3D signature motif in Hef to the $PDX_n(E/D)XK$ motif in restriction enzymes. This structural study also suggests that the XPF/Rad1/Mus81/ERCC1 proteins form a dimer through each interface of the nuclease domain and the helix-hairpin-helix domain. Simultaneous disruptions of both interfaces result in their dissociation into separate monomers, with strikingly reduced endonuclease activities.

Introduction

During the course of DNA replication, repair, and recombination processes, fork, bubble, or branched structures are produced as transient DNA intermediates with branches of duplex or single-stranded DNA. They are targeted for efficient and accurate processing by various protein factors or enzymes [1]. Recent studies have highlighted the importance of these DNA structures for the maintenance of genome stability. In fact, defects in the processing of these intermediates cause various impediments, which are connected to genetic diseases or cancer. To circumvent these problems, cells contain multiple proteins that act on these intermediates. In bacteria, the RecG helicase recognizes and processes stalled replication forks, while the RuvABC proteins act on the Holliday junction, which is the universal intermediate of homologous recombination [2–4]. However, it remains unclear how the corresponding DNA intermediates are processed in eukaryotic cells.

The XPF/Rad1/Mus81-dependent nucleases belong to one of the enzyme classes that act on the DNA intermediates with irregular structures. These proteins, which exist in various cells from yeast to human, exhibit endonuclease activities specific for DNA tertiary structures, but not nucleotide sequences. The XPF protein was identified as the catalytic component of a structure-specific endonuclease, which processes bubble structures containing damaged DNA in mammalian nucleotide excision repair [5, 6] or flap DNA during homologous recombination [7]. Together with ERCC1, this protein cleaves DNA duplexes adjacent to a 3′ single-stranded flap (Figure 1A). A similar reaction is facilitated by its yeast homologs in the Rad1-Rad10 complex [8].

Mus81 is a nuclease with sequence similarity to XPF. This protein is crucial for normal cell growth and for meiotic recombination [9–11]. Two-hybrid experiments have revealed that Mus81 interacts with the FHA domain of Cds1 checkpoint kinase in *Schizosaccharomyces pombe* [9] and the Rad54 recombination protein in *Saccharomyces cerevisiae* [10], suggesting that the protein plays some roles in recombination and checkpoint signaling. Furthermore, yeasts, lacking both Mus81 and *E. coli* RecQ homolog Sgs1/Rqh1, are synthetic lethal, highlighting its importance in genome integrity and in the replication/recombination process. The Mus81 proteins are conserved from yeast to human and share a similar branch structure-specific nuclease activity. The Mus81 complex, partially purified from yeast and human cells using the affinity tag, was shown to cleave synthetic Holliday junctions in vitro [12, 13]. On the other hand, the recombinant yeast Mus81 complex, expressed in *Escherichia coli*, preferred synthetic fork structures as cleavage substrates to the Holliday junction [14, 15]. Similar substrate preference was also observed in the partially purified human Mus81 complex [16]. Although in vivo substrates for the Mus81 complex remain unidentified, biochemical analyses indicated that its cleavage sites are similar to that for XPF/Rad1, which cleaves a DNA duplex adjacent to the 3′ branched site (Figure 1A).

The XPF/Rad1/Mus81 proteins contain the $ERKX_3D$ signature sequence [17], which is thought to be involved in metal-dependent nuclease activities, as revealed from recent mutagenesis and biochemical analyses of XPF [18]. However, the precise role of each residue in the cleavage reaction is still unclear. In addition to this signature motif, the helix-hairpin-helix DNA binding motif is also shared by all of the proteins. Both the XPF and ERCC1 proteins contain two tandemly repeated helix-hairpin-helix motifs near their carboxyl terminal (C-terminal) ends (Figure 1B). By contrast, the two helix-hairpin-helix motifs in Mus81 are separated by the insertion of a large segment, which consists of more than 400 amino acids and contains the $ERKX_3D$ signature motif [10, 13]. The amino terminal (N-terminal) regions of XPF and Rad1 include additional large domains (~500 amino acids), which are assumed to form a helicase-like domain with a displaced ATP binding site [17].

The members of the XPF/Rad1/Mus81-dependent

*Correspondence: morikawa@beri.or.jp

[3]Present address: Laboratory of Protein Chemistry and Engineering, Faculty of Agriculture, Kyushu University, Hakozaki, Higashi-ku, Fukuoka-shi, Fukuoka 812-8581, Japan.

Key words: DNA repair; structure-specific nuclease; mutational analysis; domain analysis; oligomerization; X-ray crystallography

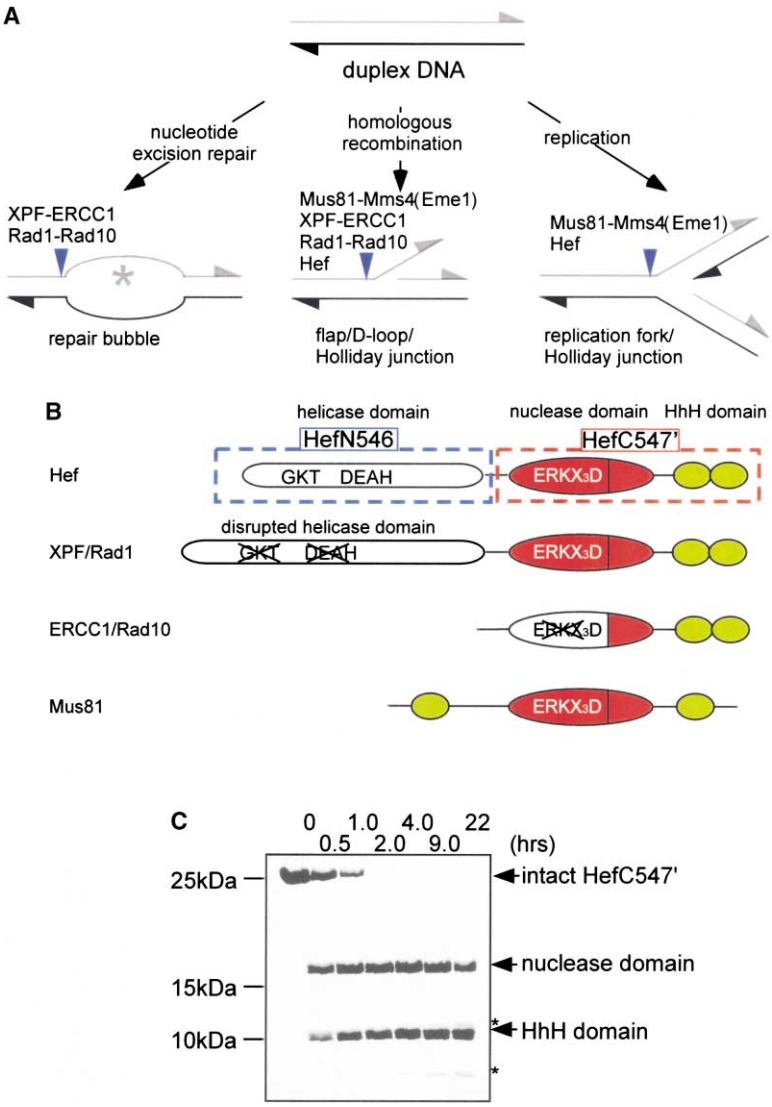


Figure 1. Domain Analysis

(A) Classification and DNA substrates of the XPF/Rad1/Mus81-dependent nucleases. Blue arrowheads indicate the cleavage sites on DNA and the asterisk denotes the damaged position.

(B) Schematic diagram of the XPF/Rad1/Mus81/ERCC1 proteins. Each domain is colored and shaped differently. HefN546 and HefC547′ are indicated by blue and red boxes, respectively.

(C) Limited proteolysis of HefC547′ by subtilisin protease. HefC547′ was mixed with subtilisin and was incubated at room temperature for the indicated amounts of time. Aliquots were analyzed by SDS-PAGE followed by Coomassie staining. Arrows on the right represent the produced HefC547′ fragments and the asterisks represent further digested products of the nuclease domain. Molecular weight markers are shown on the left.

nuclease form heterodimeric complexes, such as XPF with ERCC1 [19] and Rad1 with Rad10 [20]. Likewise, Mus81 forms a complex with Mms4 and Eme1 in *S. cerevisiae* and *S. pombe*, respectively, although the counterpart of the human Mus81 has not been identified yet [12, 14]. In both cases, the absence of their own counterparts causes the instability of the corresponding proteins, and the nuclease activity is absent [12, 14, 21–24].

Hef was initially identified from *Pyrococcus furiosus*, during screening for a protein factor that stimulates Holliday junction resolution by the Hjc resolvase [25]. The function of Hef in vivo is unknown; however, biochemical analyses revealed that Hef actually bears an endonuclease activity specific for flap or fork structures, and that this cleavage activity is independent of Hjc. The multiple alignment of homologous sequences suggests that this protein is divided into two distinct regions. The N-terminal region shows similarity to the superfamily 2 (SF2) helicase, which resembles the eukaryotic Mph1 protein involved in genome stability. The C-terminal region exhibits profound similarity to the XPF and Mus81 nucleases with the conserved $ERKX_3D$ signature motif. Most archaea contain a homolog of the XPF/Rad1/Mus81 protein [25] and possess both of these regions, but some species, such as *Sulfolobus solfataricus* and *Aeropyrum pernix*, lack the helicase region. Interestingly, the helicase motif in the archaea appears to be functional, in contrast to XPF, in which the nucleotide binding motifs are displaced by other residues [17].

To gain insights into the three-dimensional (3D) structure and function of the XPF/Rad1/Mus81-dependent nuclease, we have determined the crystal structure of the central nuclease domain of Hef. The 1.8 Å resolution structure revealed that this domain folds into architecture similar to that of restriction endonucleases, including the positions of the active site residues. The combined approach of biochemical and mutational analyses also revealed the novel domain organization and dimerization mode, which define the function of the intact molecules. These results provide important insights into the DNA cleavage and substrate recognition mecha-

nisms of this XPF/Rad1/Mus81-dependent endonuclease family.

Results

Domain Architecture of HefC547′

We failed to overproduce the full-length Hef protein (763 residues) in *E. coli*. Instead, two fragments, corresponding to the N-terminal helicase (HefN546, residues 1–546) and the C-terminal nuclease (HefC547′, residues 547–763 containing the F600L mutation) regions, were successfully produced and purified (Figure 1B), thus allowing us to characterize their biochemical properties [25]. Subsequent biochemical analyses revealed that the F600L mutation does not affect the in vitro function of HefC547′ (K.K. and Y.I., unpublished results). We found that both of these regions have preferable interaction with fork- or flap-structured DNA compared to other DNA substrates such as normal linear DNA duplexes, bubbles, or the Holliday junction [25].

HefC547′ shows substantial sequence similarity to the XPF/Rad1/Mus81 proteins, which share the $ERKX_3D$ signature sequence involved in nuclease activity [25]. The sequence comparison suggests that Hef adopts a similar domain organization to those in the XPF/Rad1/Mus81 proteins, which contain the $ERKX_3D$ sequence and the helix-hairpin-helix motifs (Figure 1B). We performed limited proteolysis to examine the structural core of the HefC547′, and found that it was so sensitive to subtilisin treatment that the intact molecule was completely digested into two major fragments within 2 hr (Figure 1C). A combination of N-terminal sequencing and mass spectrometry revealed that the cleavage took place between the nuclease domain and the following helix-hairpin-helix (HhH) domain (Figure 2A). These results suggest that the C-terminal one third of Hef is composed of the stable cores of the nuclease domain and the HhH domain.

Overall Structure of the Hef Nuclease Domain

We attempted to crystallize HefC547′ΔC4, lacking the C-terminal 4 residues, in a 150 mM NaCl solution containing a very small amount of subtilisin. This crystallization solution, concomitant with partial digestion, yielded atomic resolution crystals. Analyses of the crystal content using SDS-PAGE, mass spectroscopy, and amino acid sequencing, revealed that it consists of the nuclease domain alone and lacks the HhH domain (data not shown). The refined crystal structure of the Hef nuclease domain includes 132 amino acids, ranging from residues 550 to 681, and 106 water molecules (Table 1).

The Hef nuclease folds into a compact single-domain structure, which forms a homodimer related by a crystallographic 2-fold axis (Figure 2B, red and blue). The nuclease adopts an α/β structure, where the central β sheet is flanked by a number of α helices. The DALI server analysis [27] revealed that the folding of the Hef nuclease domain is quite similar to those of Vsr, a prokaryotic mismatch repair endonuclease, and Hjc, an archaeal Holliday junction resolvase, both adopting the type II restriction endonuclease fold, thus indicating that the Hef nuclease belongs to this restriction endonuclease family (Figure 2C). Notably, two v-shaped helical pairs ($\alpha1'$-$\alpha1$ and $\alpha2$-$\alpha3$) form a groove, which assembles the conserved residues that are rich in acidic amino acids (Figure 2D). Overall, the structural features of this acidic groove resemble the active sites of various nucleases, which belong to the restriction endonuclease family. The dimer interface, with a buried surface area of 2043 $Å^2$, is predominantly formed by the C-terminal segment (residues 630–681), which folds into an α-β-α structure with the shape of the letter N (Figure 2E). The side-by-side arrangement of this N-shaped structure allows extensive hydrophobic and electrostatic interactions in the dimer interface, such as intersubunit hydrophobic interactions between the side chains of $\alpha4$ and $\alpha5$, and those between the $\beta6$ hydrophobic side chains of each subunit.

The sequences of the nuclease segment and the dimer interface segment are well conserved among the XPF/Rad1/Mus81-dependent nuclease family members. Notably, the sequences, when compared with those of the ERCC1 family, revealed relatively high similarity in the dimer interface (Figure 2A), although the ERCC1 family lacks the conserved $ERKX_3D$ motif (Figure 2A). Interestingly, the flexible loop between $\beta5$ and $\alpha4$ lies at the edge of the dimer interface, partly protruding into the solvent, and thus this loop is specifically cleaved by prolonged subtilisin protease treatment (Figures 1C, 2A, and 2B). The lengths of this loop vary greatly among the XPF/Rad1/Mus81/ERCC1 proteins (Figure 2A). Taken together, these findings tempt us to speculate that the $ERKX_3D$ motif was replaced by different sequences in the ERCC1 family, whereas the dimer interface remained unchanged. Thus, XPF and ERCC1 are likely to share the same mode of dimer formation with the archaeal Hef protein (Figure 1B).

Metal Coordination

We have also determined the structure of Hef nuclease domain cocrystallized with either 10 mM $MnCl_2$ or $CaCl_2$ (Table 1). A strong electron density was observed at the identical position in the simulated annealed omit maps from the Mn^{2+}- and Ca^{2+}-bound structures (Figure 3A). Both of the metals are in common, coordinated to the main chain carbonyl oxygen of R594 and the two carboxyl side chains of D583 and E593, in addition to three water molecules. The nearly full occupancy of Mn^{2+} and Ca^{2+} is suggested by the peak heights of 6.0σ and 4.0σ, respectively, and by their clear octahedral coordination scheme with the chemical groups. The Hef nuclease requires Mg^{2+} or Mn^{2+}, but not Ca^{2+}, for its catalytic activity, just like the enzymes belonging to the restriction endonuclease family [25]. However, it is unclear from the present structure why Ca^{2+} inactivates the Hef endonuclease, despite its coordination scheme being essentially identical with that of Mn^{2+}. Various structural analyses of type II restriction endonuclease crystals demonstrated that, in each case, the metals are coordinated with the chemical groups corresponding to the carbonyl group of R594, while further metal coordination has been proposed for specific DNA substrate recognition (reviewed in [28]). Two other metal-coordinated acidic residues, D583 and

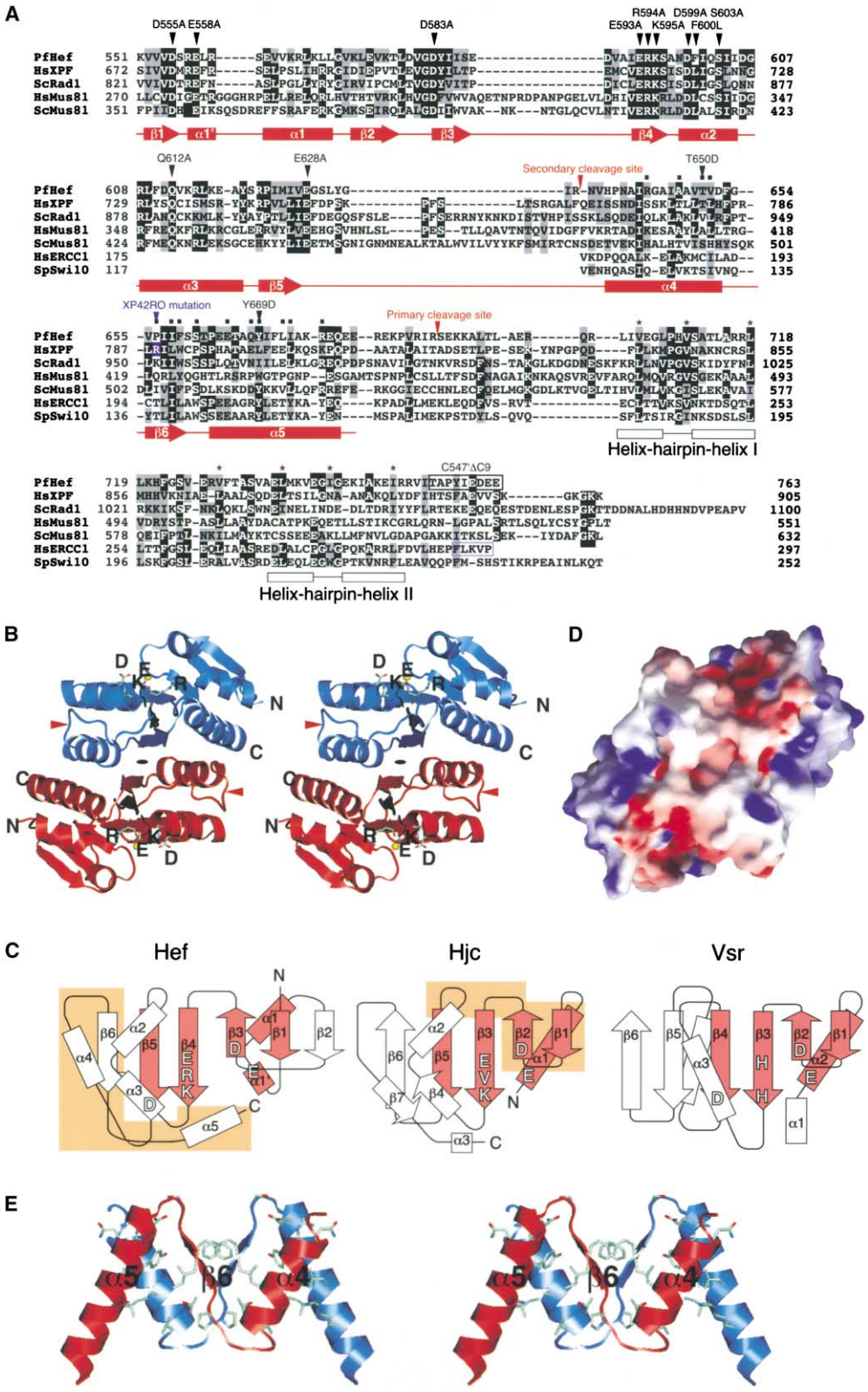


Figure 2. Structure of the Hef Nuclease Domain

(A) Sequence alignment of XPF/Rad1/Mus81 and ERCC1/Swi10. Black boxes represent residues that are identical among more than three proteins. Gray boxes represent similar residues. Secondary structure elements are indicated under the primary sequences. Cylinders, α helices; arrows, β strands. Black squares above the sequence indicate residues involved in dimerization. Protease-sensitive sites are indicated by red arrowheads, and mutation sites are shown by black arrowheads. The blue arrowhead indicates XPF mutation site and the blue box

Table 1. Data Collection and Refinement Statistics

Data Collection							
Data Set	Native	Se-Met L1 Remote	Se-Met L2 Peak	Se-Met L3 Edge	Se-Met L4 Remote	Native + $MnCl_2$	Native + $CaCl_2$
Wavelength (Å)	1.0000	0.9700	0.9791	0.9794	1.0000	1.0000	1.0000
Resolution (Å)	1.78	1.8	1.8	1.8	1.8	1.78	1.78
Completeness (%)[a]	98.8 (94.1)	99.6 (100)	100 (100)	99.6 (100)	97.7 (100)	99.2 (95.1)	98.9 (89.1)
R_{merge}[a,b]	0.058 (0.173)	0.055 (0.181)	0.056 (0.187)	0.051 (0.178)	0.052 (0.118)	0.060 (0.200)	0.052 (0.246)
f′/f″	—	−4.06/3.76	−7.66/3.84	−9.80/3.84	−3.47/0.51	—	—
Unique reflections	15,983	15,786	15,845	15,783	15,484	15,973	16,139

Refinement Statistics				
	Native	+$MnCl_2$	+$CaCl_2$	Se-Met
Resolution (Å)	50.0–1.78	50.0–1.78	50.0–1.78	50.0–1.8
Protein atoms (avg. B value)	1030 (23.2)	1045 (23.1)	1045 (23.3)	1030 (22.3)
Solvent molecules (avg. B value)	106 (33.5)	84 (33.1)	94 (31.9)	105 (34.9)
Ion molecule (B value)	—	1 (32.6)	1 (33.0)	—
R factor/free R factor (%)	22.4/24.8	22.8/25.2	21.5/24.6	22.3/23.8
Rms bond lengths (Å)	0.004	0.005	0.005	0.005
Rms bond angles (°)	1.05	1.09	1.08	1.10
Ramachandran plot				
Most favored	94.9%	94.1%	94.1%	94.9%
Allowed	5.1%	5.9%	5.9%	5.1%

[a] Numbers in parentheses represent statistics in the highest resolution shell.
[b] R factor = $\Sigma|F_{obs} - F_{calc}|/\Sigma f_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. The R_{free} was calculated using a randomly selected 10% of the data set that was omitted through all stages of refinement.

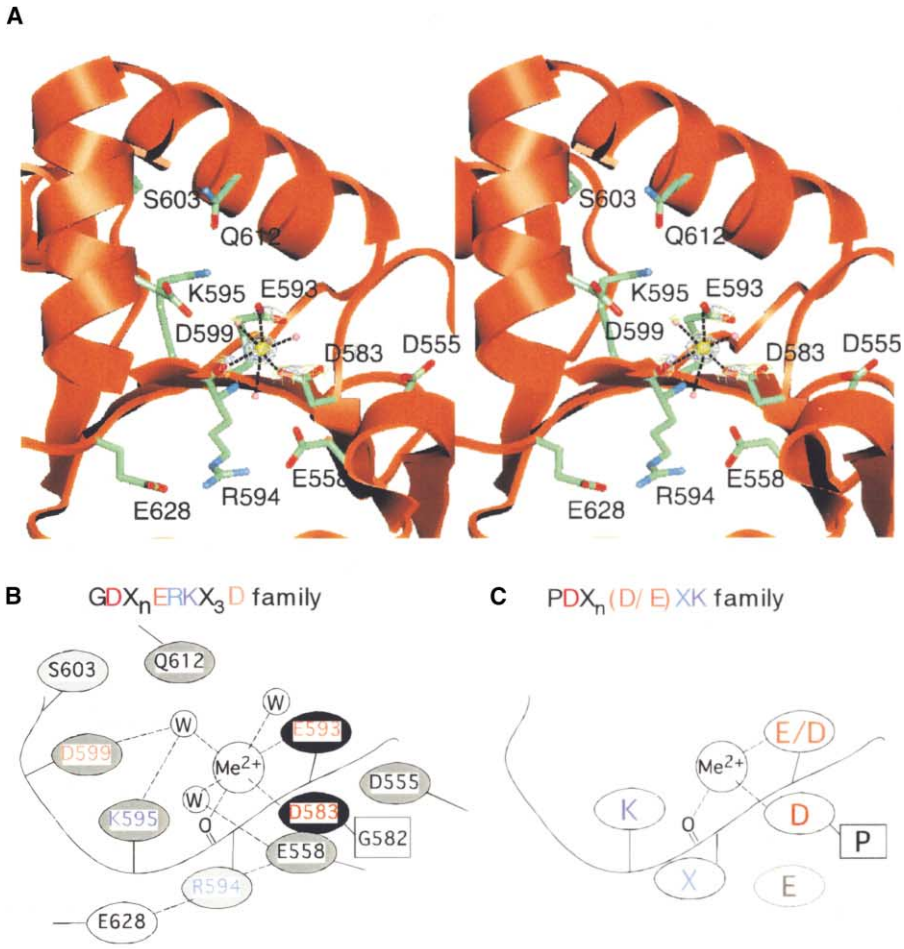
E593, in the nuclease motif of Hef protein also appear to be equivalent with the 2 acidic residues in the canonical motif of the type II restriction endonucleases. This similarity in the metal coordination between the restriction endonucleases and Hef may reflect a cognate mechanism in catalysis.

Structural Similarity in the Catalytic Centers of Hef and Restriction Endonucleases

To address whether the residues in the $ERKX_3D$ motif are actually involved in the cleavage activity of Hef endonuclease, we carried out alanine scanning mutational analyses (Figures 3B and 3D). The D583A and E593A mutations completely abolished the endonuclease activity of HefC547′. This could be due to the disruption of the metal coordination scheme essential for the catalytic reaction. In addition, D555A, E558A, K595A, and D599A mutations severely impaired the activity, while R594A and Q612A mutations showed partial defects in the activity. On the other hand, the replacements of Ser603 and Glu628 by alanine did not cause serious effects. These results are consistent with the general argument that substitutions for polar residues, involved in the conformational maintenance of the catalytic center, cause serious impairments of the nuclease activity.

The type II restriction endonucleases share the $PDX_n(D/E)XK$ motif, which is crucial for catalytic activity [28]. The 2 acidic residues within this motif coordinate divalent cations, and the lysine residue is most likely to activate the attacking water (Figure 3C). The structural and sequence comparisons of Hef/XPF/Mus81 with the restriction endonuclease families suggest that the conserved $ERKX_3D$ signature motif could be expanded on the N-terminal side and included within a new motif, GDX_nERKX_3D, which coincides better with the $PDX_n(D/E)XK$ motif in restriction endonucleases (Figure 3B). In this signature motif, all of the polar residues are involved in either the coordination of the metal or the hydrogen bonding with the coordinated water molecules. The major difference in the motif sequences between the two families is that the hydrophobic X residue, within the "(D/E)XK" motif of the restriction endonuclease family, is replaced by arginine in the XPF/Rad1/Mus81 proteins (Figure 2A). The basic side chain of Arg594 in HefC547′

indicates ERCC1 truncated mutation. The conserved hydrophobic residues in helix-hairpin-helix motif are indicated by the asterisks. The first helix-hairpin-helix motif in the Mus81 protein is located at the N terminus and thus the sequence is not well aligned for Mus81 in the helix-hairpin-helix II region. Pf, *Pyrococcus furisosus*; Hs, *Homo sapiens*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*.
(B) Stereo diagram of the Hef nuclease domain crystal structure. Two molecules, related by the crystallographic 2-fold symmetry, are colored in red and blue. The $ERKX_3D$ signature motif is shown by amino acid side chains of its constituents and the bound metal is shown as a yellow sphere. The red arrowhead indicates the subtilisin-sensitive site. The 2-fold axis runs through the middle, between the two molecules.
(C) Topology comparison of the Hef nuclease domain with Hjc and Vsr. The α helices are shown as rectangles and the β strands as arrows. Conserved secondary structures are colored red and the potential amino acid residues involved in catalysis are shown as single letter representations. Regions involved in dimerization contacts are colored orange.
(D) Surface potential of the Hef nuclease domain. A GRASP surface [44], viewed in the same direction as in (B), is colored by the electrostatic potential (−10 kT/e^- to 10 kT/e^-). Positive regions are colored blue and negative regions are colored red.
(E) Stereo view of the nuclease dimer interface. Side chains involved in intersubunit interactions are shown in licorice models.


A
S603
Q612
K595
E593
D599
D583
D555
E628
R594
E558
S603
Q612
K595
E593
D599
D583
D555
E628
R594
E558
B
GDXnERKX3D family
S603
Q612
W
W
D599
E593
Me2+
D555
W
K595
D583
G582
O
E558
R594
E628
C
PDXn(D/E)XK family
E/D
Me2+
K
D
P
O
X
E
D
WT
E558A
E593A
K595A
S603A
E628A
D555A
D583A
R594A
D599A
Q612A
C
substrate
product

interacts with the carboxyl groups of Glu558 and Glu628 (Figure 3A), and these electrostatic interactions are conserved within the XPF/Rad1/Mus81 proteins. Mutation of either Arg594 or Glu558 impairs the nuclease activity of HefC547′, underscoring the importance of their interaction. The position of Glu558 in HefC547′ appears to correspond to Glu9 of Hjc, implying their similar roles in catalysis (Figures 2C, 3B, and 3C). In addition, the members of the restriction endonuclease family lack corresponding residues to both Asp599 within the signature motif and the Gln612 conserved in XPF/Rad1/Mus81 proteins, which lies outside the motif (Figures 2A, 3A, and 3B). These specific residues for the XPF/Rad1/Mus81-dependent nuclease may be involved in the recognition of branched DNA.

Dimer Interfaces

All of the nucleases that belong to the XPF/Rad1/Mus81-dependent nuclease family form dimers. When single subunits of these dimeric proteins alone were produced without their counterparts, extremely low amounts were obtained as the soluble protein, and their biochemical properties were generally unstable, implying that the dimer formation stabilizes the proteins as the active form in cells [12, 14, 20–24]. Thus, we examined the oligomeric states of HefC547′ by size exclusion chromatography (Figure 4A). The full-length HefC547′ protein eluted as a single peak with a 44 kDa molecular mass, which corresponded to the homodimer. This tendency of dimer formation is similar to those of XPF and Rad1, which form heterodimers with the ERCC1 and Rad10 proteins, respectively. The same line of experiments, performed on fragments from a partially digested HefC547′ protein, indicated that the nuclease domain and the HhH domain elute as discrete bands with molecular masses of 25 kDa and 21 kDa, respectively (Figure 4B). This result suggests that the HhH domain also forms a dimer, like the adjacent nuclease domain, and that these two domains are completely independent of each other within each HefC547′ subunit.

To gain more detailed insights into the dimerization of this nuclease family, we introduced several mutations into HefC547′ and investigated their influence upon dimer formation by a gel filtration analysis, in comparison with the crystal structure of the nuclease domain. In fact, the replacement of Thr650 in the dimeric interface by asparagine abolishes the dimer formation by the nuclease domain alone. However, the full-length HefC547′ containing the T650D mutation can still form a dimer in solution (Figure 4C). On the other hand, the nuclease domain of HefC547′T650D produced by subtilisin treatment showed a shift in its elution profile, whereas the HhH domain peak remained unchanged after the same treatment (Figure 4D). Taken together, the results indicate that the T650D mutation disrupts the dimer interface of only the nuclease domain, independently of the HhH domain of HefC547′ protein. A similar result was observed with the Y669D mutation, which presumably would disturb the dimer interface (data not shown).

As for the HhH domain of the XPF/Rad1/Mus81-dependent nucleases, the C-terminal region of ERCC1 is known to participate in the heterodimerization with XPF [29]. In fact, the deletion of the C-terminal 5 residues from ERCC1 protein impairs the interaction with XPF, suggesting that this terminal region partly constitutes the dimer interface in the HhH domain. For comparison, we created a similar mutant HefC547′ protein, in which the C-terminal 9 residues were deleted (HefC547′ΔC9; Figure 2A). Gel filtration analyses showed that the full-length HefC547′ΔC9 forms the dimer in solution (Figure 4E). However, truncation of the C-terminal 9 residues disrupts the dimer formation of the HhH domain, but not of the nuclease domain in HefC547′ (Figure 4F).

If our assumption that HefC547′ forms the dimer through a combination of the two interfaces is true, then dual mutations should produce monomeric species of HefC547′. When the full-length HefC547′ with mutations in each of the two domains (HefC547′T650D/ΔC9) was analyzed by gel filtration, they eluted as monomers, as expected (Figure 4G). We also confirmed that the individual domains of HefC547′T650D/ΔC9, produced by subtilisin treatment, are monomeric in solution (Figure 4H). Thus, the HefC547′ molecule contains two dimer interfaces, which function independently with each other.

Dimer Formation Is Essential for the Cleavage Activity of Fork DNA

The enzymatic activity of the XPF/Rad1/Mus81-dependent nuclease is tightly coupled with heterodimer formation. The absence of the counterpart protein not only destabilizes the protein against proteolytic cleavage or denaturation, but also inactivates the nuclease [12, 14, 20–24]. We used the dimer mutant HefC547′ proteins to analyze the relationships between dimer formation and synthetic fork-DNA binding and cleavage.

First, we analyzed the nonspecific DNA binding of the proteins, using a synthetic 50-mer DNA duplex. Under the condition that abolishes the complex formation of

Figure 3. Active Site Analyses of the Hef Nuclease Domain

(A) The metal binding site of the Hef nuclease is shown in a stereo diagram. Conserved side chains in the vicinity of the metal binding site are shown in licorice models. The divalent metal and the coordinating water molecules are shown as yellow and gray balls. The electron densities of Ca^{2+} (green) and Mn^{2+} (blue) after the simulated annealed omit map are contoured in 2 σ and 3 σ, respectively.

(B) Schematic diagram of the Hef nuclease active site, containing the GDX_nERKX_3D motif. Side chains are presented as ovals, and their contribution to the activity is colored in a gray scale. Critical residues are colored black.

(C) Schematic diagram of the $PDX_n(E/D)XK$ motif found in restriction endonucleases. The corresponding amino acids between the GDX_nERKX_3D motif and $PDX_n(E/D)XK$ motif are colored similarly.

(D) Mutation analysis of active site residues and conserved residues. ^{32}P-labeled synthetic fork substrate (200 fmol) was mixed with 2 pmol and 20 pmol of wild-type or mutant HefC547′ proteins at 60°C for 30 min. Products were analyzed by denaturing PAGE and were detected by autoradiography.

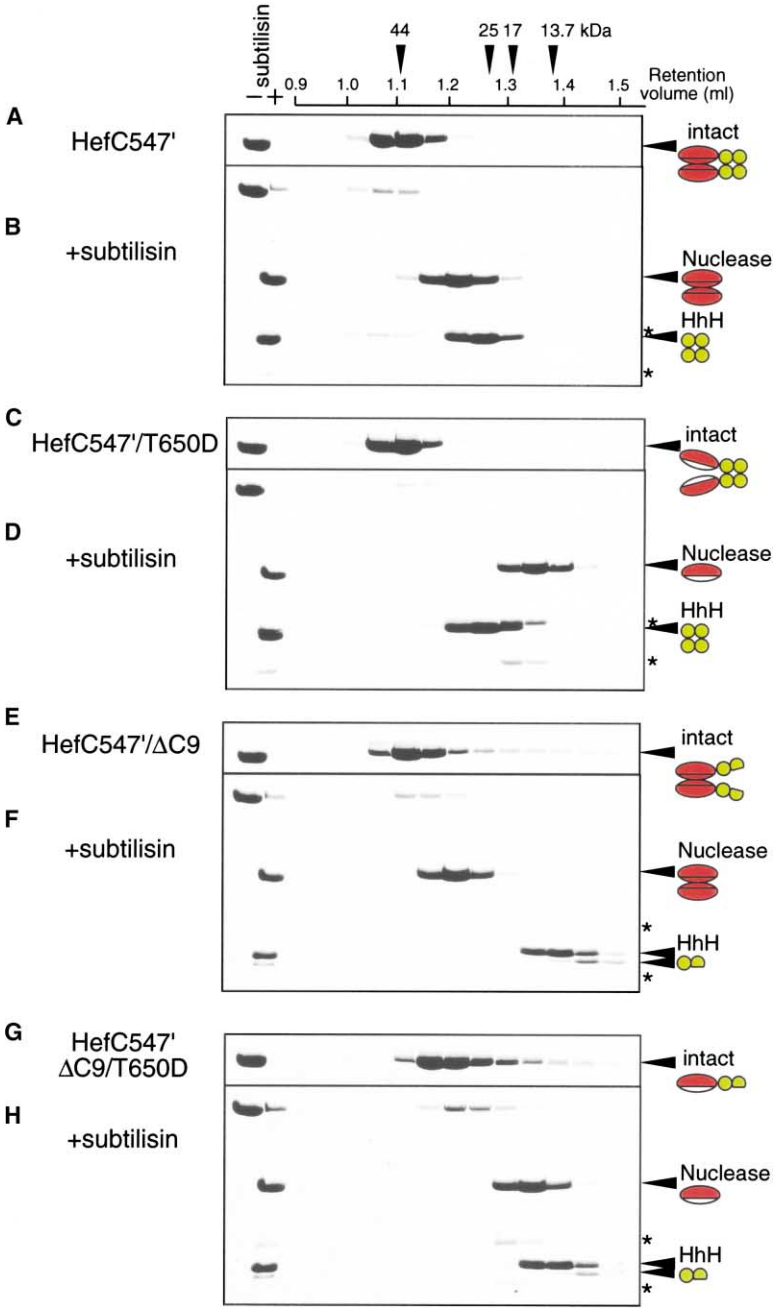


Figure 4. Analysis of Dimer Interfaces

Gel filtration analysis of intact HefC547′ and its proteolyzed fragments.
(A) Intact HefC547′.
(B) Proteolyzed HefC547′.
(C) HefC547′T650D.
(D) Proteolyzed HefC547′T650D.
(E) HefC547′ΔC9.
(F) Proteolyzed HefC547′ΔC9.
(G) HefC547′T650D/ΔC9.
(H) Proteolyzed HefC547′T650D/ΔC9.
Protein (250 pmol) was loaded onto a Superdex G75 column, analyzed by SDS-PAGE, and stained by Coomassie. Arrows and asterisks represent bands of HefC547′ fragments, as in Figure 1C. Schematic drawings are shown in the right column for the mutation and the oligomerization state. Coloring schemes are the same as in Figure 1B. The proteins containing truncation mutations were subjected to further cleavage in the helix-hairpin-helix domain, and two bands appeared after proteolysis.

the wild-type protein (Figure 5A, lane 3), the dimer mutant HefC547′ proteins formed a protein-DNA aggregate that could not get into the gel. This effect was observed in the protein with a mutation in either the nuclease domain (HefC547′T650D; Figure 5A, lane 5) or the HhH domain (HefC547′ΔC9; Figure 5A, lane 7), whereas the two HhH mutants (HefC547′ΔC9 and HefC547′T650DΔC9; Figure 5A, lanes 7 and 9) exhibited the increase of aggregation. Next, we analyzed the fork structure-specific binding activity in the presence of competitor DNA, double-stranded (ds) poly-dI-dC. Under this condition, wild-type HefC547′ formed a specific protein-DNA complex, which was observed as a retarded band on the gel (Figure 5B, lane 3). HefC547′T650D with a mutation in the nuclease domain formed a similar protein-DNA complex (Figure 5B, lane 5). The two HhH mutants (HefC547′ΔC9 and HefC547′T650DΔC9) also exhibited a similar band shift, while they additionally formed a protein-DNA aggregate similar to those observed in the case of 50-mer dsDNA (Figure 5B, lanes 7 and 9). This aggregate disappeared by increasing the competitor DNA (data not shown). These results indicate that all of the proteins with mutations in each of the dimer interfaces could recognize synthetic fork substrate, although their specificity to the fork structure is reduced by the mutations.

Next, we analyzed the cleavage activity of the synthetic fork DNA by the wild-type and dimer mutant HefC547′. Under the condition that allows the wild-type

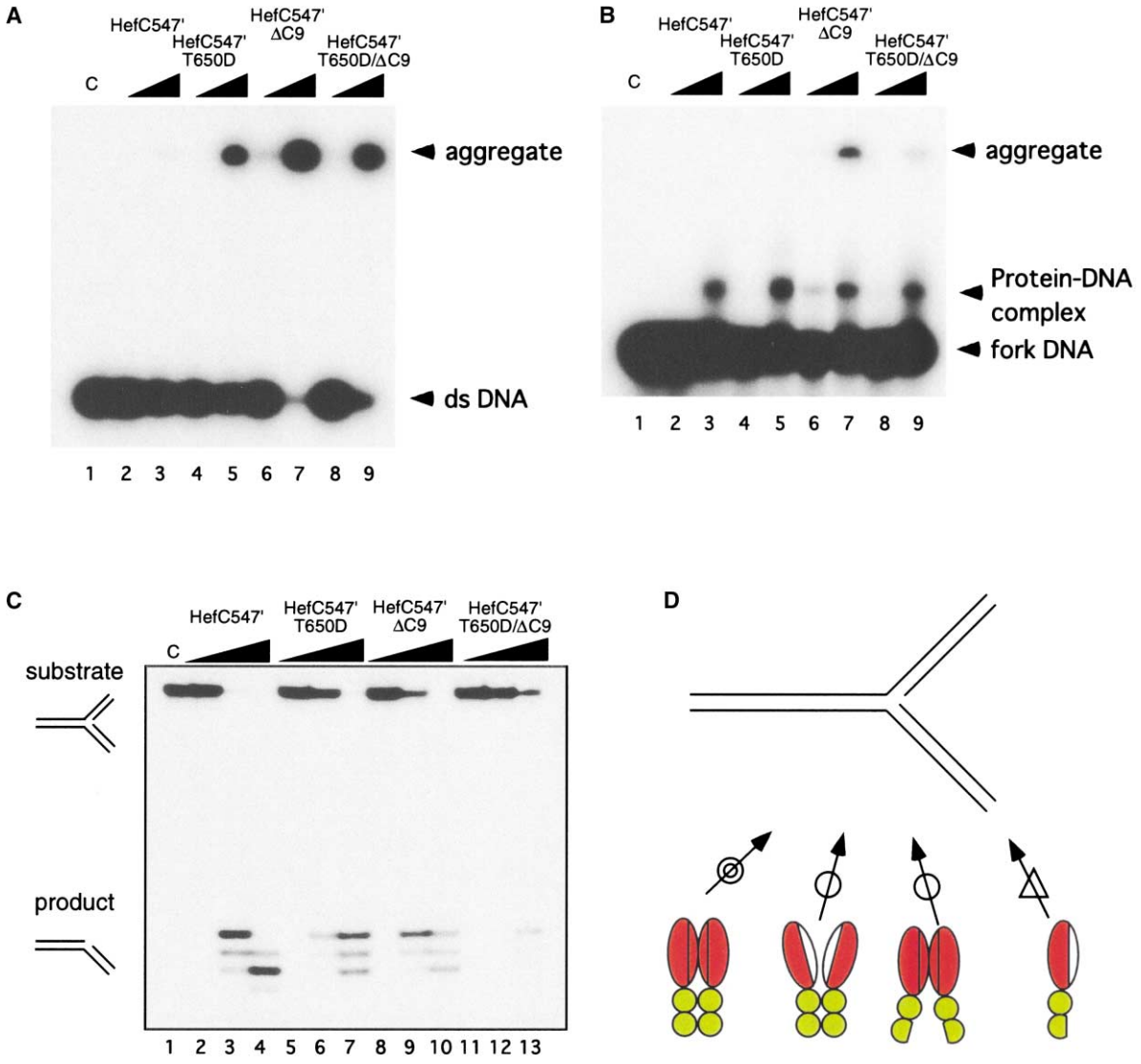


Figure 5. Endonuclease Activity of Dimer Mutants

(A) DNA binding activity of dimer mutants to 50-mer dsDNA. ^{32}P-labeled 50-mer dsDNA (200 fmol) was mixed with 500 fmol or 5 pmol of wild-type or mutant HefC547′ proteins at 37°C for 10 min. Protein-DNA complexes were crosslinked with glutaraldehyde. Products were analyzed on native PAGE by autoradiography.

(B) DNA binding activity of dimer mutants to the synthetic fork substrate. The assay was carried out as in (A) using 200 fmol of ^{32}P-labeled synthetic fork substrate in the presence of 0.15 ng ds poly-dI-dC.

(C) Nuclease activity of dimer mutants. Synthetic fork substrate (200 fmol) was mixed with 20 fmol, 200 fmol, and 2 pmol of wild-type or mutant HefC547′ proteins at 50°C for 60 min. Products were analyzed by denaturing PAGE and were detected by autoradiography.

(D) Schematic drawing of dimerization and endonuclease activity of HefC547′ to the fork substrate.

to cleave the substrate (Figure 5C, lanes 2–4), both the HefC547′T650D and HefC547′ΔC9 mutants showed substantial nuclease activity, with several fold reduction in comparison with the intact protein (Figure 5C, lanes 5–7 and lanes 8–10). By contrast, the dual mutations in both the nuclease and HhH domains reduced the activity by more than 100-fold (Figure 5C, lanes 11–13). Our crystal structure indicates that neither the HefC547′ T650D mutation nor the HefC547′ΔC9 mutation disrupts the catalytic center of the nuclease domain, and hence we believe that this striking reduction was caused by the incapability of the mutants to form the dimer. Taken together, we conclude that simultaneous dimer formation in both the nuclease and HhH domains is crucial for HefC547′ to recognize the fork DNA correctly (Figure 5D).

Discussion

Relationships between the XPF/Rad1/Mus81-Dependent Nuclease Family and the Restriction Endonuclease Family

Restriction endonucleases are constituents of the restriction modification system found ubiquitously in prokaryotes, which functions mainly to protect the host genome from foreign DNA. This system is classified into

three groups with different subunit organizations. The type II restriction endonuclease contains a small endonuclease subunit with an approximate molecular mass of 30 kDa. On the other hand, the type I and III restriction endonucleases contain a larger endonuclease subunit (~100 kDa), and the helicase domain is fused to the nuclease domain just like the intact Hef protein. Although all restriction endonucleases contain conserved nuclease domains, their sequence conservation is limited to the active site residues. Nonetheless, the 3D structures of each nuclease domain revealed the common structural fold, which is composed of a five-stranded β sheet flanked by α helices.

Recent structural studies have revealed that several nucleases, such as the G-T mismatch recognition Vsr endonuclease from *E. coli* [30] and the Holliday junction-specific Hjc endonuclease from archaea [31, 32], also fold into architectures similar to that of the restriction endonuclease. In particular, the latter enzyme shares the same active site residues as those in the restriction endonucleases [33–35]. On the other hand, it has been believed that the XPF/Rad1/Mus81-dependent nucleases are distinct from the restriction endonuclease family in terms of their primary structures, because of the specific signature sequence of the former family [17]. However, our X-ray structural and biochemical analyses of Hef endonuclease have revealed interesting relationships between the two families, in terms of molecular evolution: the nuclease domain of the XPF/Rad1/Mus81 proteins shares a common origin with the restriction endonuclease family in prokaryotes. The signature sequence of XPF/Rad1/Mus81 proteins is closely related to the consensus sequence essential for the catalytic activity of restriction endonucleases. Particularly, the N-terminal moiety of the GDX_nERKX_3D sequence motif in the XPF/Rad1/Mus81 proteins corresponds to the $PDX_n(E/D)XK$ motif in restriction endonucleases.

Domain Organization of the XPF/Rad1/Mus81-Dependent Nucleases

It has been shown by biochemical analyses that Hef consists of the helicase and nuclease regions, both of which favor the fork structure as substrate [25]. Our present study reveals that the nuclease region is further divided into the compact nuclease domain and the HhH domain. It was postulated that the N-terminal 650 residues in XPF form the helicase fold with displaced nucleotide binding motifs, and that this fold is followed by the nuclease domain and the HhH domain (Figure 1B) [17]. ERCC1, lacking the helicase region, shows the sequence similarity in its C-terminal region to XPF [36], whereas the active site residues for the endonuclease are replaced by different residues. This distinct feature appears to be related to the unique functions of ERCC1, which interacts with XPA involved in nucleotide excision repair [37, 38]. Presumably, ERCC1 would have replaced most of the residues in the nuclease domain by more suitable residues, which contribute to the recognition of other protein partners. Interestingly, Mus81 contains two HhH motifs, which are separated by a nuclease domain [10, 13]. Because Mus81 recognizes substrates similar to that for Hef, the two HhH motifs in Mus81 may assemble to form a compact domain structure.

Strong Coupling between Dimerization and Activity in the XPF/Rad1/Mus81 Family Nucleases

Interaction mapping by immunoprecipitation revealed that human XPF and ERCC1 forms a dimer through the interface at the C-terminal region [29]. The truncated mutation in ERCC1 protein lacking five amino acids from the C terminus is unable to interact with XPF. On the other hand, the N terminus of ERCC1 is less important in the association, because more than 220 amino acids can be deleted without seriously impairing the association [29]. Similarly, in XPF, more than 800 amino acids from the N terminus can be deleted without disturbing the association with ERCC1, whereas the deletion of C-terminal 25 amino acids abolished their interaction [29]. These findings are explained well by our dimer model of HefC547′ using the HhH domain.

Intriguingly, XPF patients XP42RO carry the homozygous XPF gene with the R788W mutation, which exhibits the XPF phenotype. In vivo analyses have revealed that this mutation reduces the nucleotide excision repair activity by 5-fold in cultured cells [39]. The mutant XPF protein appears to be unstable, as reflected by its low production. This mutation site is mapped at the middle position of the β5 strand involved in dimerization of the nuclease domain (Figure 2A), and hence we presume that the R788W phenotype results from a disruption of the interface for interaction with ERCC1.

The Hef protein is a homodimer, which contains two copies of the nuclease domain and two copies of the tandem HhH domain. On the other hand, the XPF-ERCC1 complex possesses only one active nuclease domain from XPF and two HhH domains from both subunits, respectively. It has been proposed that the XPF and ERCC1 proteins may have evolved by an ancient duplication, and thus Hef is a possible candidate for the ancestor [26]. In addition, Hef and XPF-ERCC1 cleave DNA strands at a similar position in the branched structure. Therefore, it is likely that only one active nuclease domain of Hef is required for the activity, while the two copies of the HhH domains appear to participate in DNA binding. Alternatively, Hef may recognize other unknown DNA structures, which would be cleaved by both active sites. Further biochemical and structural analyses of the protein-DNA complex will illuminate the structure-specific recognition by this protein family. Nevertheless, a tentative docking examination suggests that the junction region of the fork-structured DNA is bound to the nuclease domain, and the branching DNA duplexes are recognized by the HhH domains.

Biological Implications

Our structural and biochemical analyses of the archaeal XPF/Rad1/Mus81 nuclease have clarified its domain organization and provided a detailed 3D structural view of the endonuclease domain, including the catalytic center containing the conserved GDX_nERKX_3D sequence, which are expanded from the previously proposed signature motif for the nucleases of this family. This is the first report revealing that this motif, which has been specific for bacterial restriction enzymes, is present in

eukaryotic and archaeal proteins, implying a possible origin of this nuclease family. We have shown by biochemical analyses of various mutant HefC547′ proteins that the nuclease domain and the HhH domain are present as completely separate building blocks, connected only through the flexible linker. The crystal structure of the nuclease domain revealed the structural details of its dimer interface, which involves extensive hydrophobic and polar interactions contributed by conserved residues among the XPF/Rad1/Mus81-dependent nucleases. Notably, the primary structure of the interface region is also conserved in the ERCC1 family, which forms heterodimers with XPF family proteins. The independent dimer interface is also found in the HhH domain, which consists of a 60 residue segment at the C terminus. The XPF/Rad1/Mus81-dependent nucleases recognize tertiary structures, such as the branched and bubbled structures of DNA duplexes, but not specific DNA sequences. The present study revealed that this structural specificity and the DNA cleavage activity are tightly coupled with dimerization.

Experimental Procedures

Limited Proteolysis and Gel Filtration

The HefC547′ protein (50 μM) was mixed with subtilisin (100 ng/ml) in 10 mM HEPES-NaOH (pH 7.4), 150 mM NaCl, and 1 mM $CaCl_2$. The mixture was incubated at room temperature for 16 hr. Samples were loaded onto a gel filtration column (Superdex G75 PC, Amersham Biosciences) equilibrated with buffer (50 mM Tris-HCl [pH 8.0], 0.1 mM EDTA, 0.5 mM DTT, 10% glycerol, and 0.3 M NaCl) and the fractions were analyzed by SDS-PAGE.

Preparation and Crystallization of HefC547′

HefC547′ΔC4 was constructed by creating a stop codon within HefC547′, removing four amino acids from the C terminus. Sequencing of the expression construct revealed an additional PCR-induced mutation (R608G), which is not present in the intact HefC547′ protein, although this mutation had no effect on the nuclease activity (data not shown). The protein was concentrated to 60 mg ml^{-1} by ultrafiltration (Millipore). HefC547′ΔC4 was crystallized at 20°C by the microdialysis method. Good quality crystals were obtained when 10 μl of the HefC547′ΔC4 solution was mixed with 0.3 μl of 10 μg/ml subtilisin and dialyzed in a buffer containing 100 mM tricine-NaOH (pH 8.0), 300 mM NaCl, and 0.025% NaN_3. One day later, the mother liquor was diluted with H_2O to make a 200 mM NaCl solution. Hexagonal-shaped crystals appeared within 1 day. The divalent metal-bound protein crystals were grown in solutions containing 10 mM $MnCl_2$ and $CaCl_2$.

Data Collection and Phasing

Crystals were harvested in a solution buffer containing 100 mM HEPES-NaOH (pH 7.4) and 150 mM NaCl. Crystals were cryoprotected using 1,5-pentanediol and X-ray diffraction data were collected at 100 K using a nitrogen stream. The selenomethionine data were collected at BL38B2 (SPring8) using four different wavelengths (L1 = 1.0000 Å, L2 = 0.9791 Å, L3 = 0.9794 Å, and L4 = 0.9700 Å). These data were processed by the HKL suite [40]. The crystal belongs to the space group $P3_121$, with unit cell dimensions a = b = 69.8 Å, and c = 58.5 Å $\alpha = \beta = 90^{\circ}$ and $\gamma = 120^{\circ}$, and contains one molecule per asymmetric unit. The scaling statistics are indicated in Table 1. Native and cocrystal data were also collected at BL38B2 (SPring8) using a single wavelength (1 Å). The initial phases were calculated by the MAD data using the selenomethionine data, in which a single selenium site was found by SOLVE [41]. The heavy atom parameters were refined by SHARP with SOLOMON [42]. The initial model was built using the program QUANTA (MSI). All data between 20 and 1.8 Å resolution were included during CNS refinement [43]. Ten percent of the reflections were kept separate to monitor R_{free}, and were not included in refinement. The final refinement statistics are shown in Table 1. The current model contains a protein region including residues 550–681 and 106 water molecules. Data from the cocrystal with $MnCl_2$ and $CaCl_2$ were also collected on beamline BL38B2 in SPring8. The final structure of the selenomethionine protein was used as an initial model, which was further refined by CNS. Scaling statistics and final refinement statistics are shown in Table 1.

DNA Substrate

A synthetic replication fork structure (FI) was prepared for the nuclease activity experiments. The oligonucleotides used to make fork DNA structure were described in the previous paper [25]. Synthetic 50-mer DNA duplex was prepared by annealing 5′-CCATGCC TGCACGAATTAAGCAATTCGTAATCATGGTCATAGCTGACTAC-3′ and 5′-GTAGTCAGCTATGACCATGATTACGAATTGCTTAATTCGT GCAGGCATGG-3′.

Site-Specific Mutagenesis

PCR-mediated mutagenesis was carried out using the Quick Change site-directed mutagenesis kit (Stratagene). The sequence of the mutants was verified by an ABI 3700 sequencer (Perkin Elmer).

Purification of Wild-Type and Mutant HefC547′ Proteins

Gene expression and purification of mutant proteins were performed essentially under the same conditions as described previously [25] except that HefC547′ΔC9 and HefC547′T650D/ΔC9 mutants were prepared without heat treatment. SP Sepharose (Amersham Biosciences), heparin (Amersham Biosciences), and hydroxylapatite (Bio-Rad) columns were used for purification.

Gel Mobility Shift Assay

HefC547′ was incubated at 37°C for 10 min in 10 μl of binding buffer (10 mM triethanolamine [pH 7.5], 1 mM EDTA, and 0.1 mg/ml BSA) with 20 nM 5′-^{32}P-labeled fork substrate (FI) or 20 nM 5′-^{32}P-labeled 50-mer double-stranded DNA. For the reaction with the fork substrate, 0.5 μl of 0.3 μg/μl ds poly-dI-dC DNA was added as a competitor DNA. The protein-DNA complex was crosslinked with 0.4 μl 10% glutaraldehyde and incubated at room temperature for 5 min. The reaction was terminated by the addition of 5 μl loading buffer containing 20 mM Tris-acetate (pH 8.0), 10% glycerol, and 0.1% bromophenol blue. Aliquots were analyzed by 6% PAGE in TAE buffer at 15 mA for 50 min. The gel was dried and the bands were visualized by autoradiography.

Cleavage of Synthetic Fork Substrate by HefC547′

HefC547′ was incubated at 60°C for 30 min (active site mutants) or 50°C for 60 min (dimer mutants) in 10 μl of cleavage buffer (10 mM Tris-HCl [pH 8.8], 10 mM $MgCl_2$, 1 mM DTT, and 0.1 mg/ml BSA) with 20 nM 5′-^{32}P-labeled fork substrate (FI). The reaction was terminated by the addition of 10 μl formamide containing 0.1% xylene cyanol. Aliquots were analyzed by 15% denaturing PAGE followed by autoradiography.

Acknowledgments

The authors would like to thank Ryosuke Fujikane for the initial preparation of the HefC547′ protein, and Daisuke Tsuchiya for help in data collection and structural determination. We also acknowledge Kazuya Osaka and Hiroyuki Toh for model building and sequence analysis. We are grateful to Naoto Yagi and Keiko Miura for their help in the use of SPring8 BL38B2. T.N. is a research fellow of the Japan Society for the Promotion of Science. This research was partly supported by NEDO (New Energy and Industrial Technology Development Organization).

Received: October 9, 2002
Revised: January 8, 2003
Accepted: January 24, 2003
Published: April 1, 2003

References

1. Friedberg, E., Walker, G., and Siede, W. (1995). DNA Repair and Mutagenesis (Washington, DC: ASM press).
2. Kowalczykowski, S.C. (2000). Initiation of genetic recombination and recombination-dependent replication. Trends Biochem. Sci. *25*, 156–165.
3. Michel, B., Flores, M.J., Viguera, E., Grompone, G., Seigneur, M., and Bidnenko, V. (2001). Rescue of arrested replication forks by homologous recombination. Proc. Natl. Acad. Sci. USA *98*, 8181–8188.
4. Cox, M.M. (2001). Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. Annu. Rev. Genet. *35*, 53–82.
5. Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859–868.
6. Park, C.H., and Sancar, A. (1994). Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins. Proc. Natl. Acad. Sci. USA *91*, 5017–5021.
7. de Laat, W.L., Appeldoorn, E., Jaspers, N.G., and Hoeijmakers, J.H. (1998). DNA structural elements required for ERCC1-XPF endonuclease activity. J. Biol. Chem. *273*, 7835–7842.
8. Bardwell, A.J., Bardwell, L., Tomkinson, A.E., and Friedberg, E.C. (1994). Specific cleavage of model recombination and repair intermediates by the yeast Rad1-Rad10 DNA endonuclease. Science *265*, 2082–2085.
9. Boddy, M.N., Lopez-Girona, A., Shanahan, P., Interthal, H., Heyer, W.D., and Russell, P. (2000). Damage tolerance protein Mus81 associates with the FHA1 domain of checkpoint kinase Cds1. Mol. Cell. Biol. *20*, 8758–8766.
10. Interthal, H., and Heyer, W.D. (2000). MUS81 encodes a novel helix-hairpin-helix protein involved in the response to UV- and methylation-induced DNA damage in *Saccharomyces cerevisiae*. Mol. Gen. Genet. *263*, 812–827.
11. Mullen, J.R., Kaliraman, V., Ibrahim, S.S., and Brill, S.J. (2001). Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. Genetics *157*, 103–118.
12. Boddy, M.N., Gaillard, P.H., McDonald, W.H., Shanahan, P., Yates, J.R., III, and Russell, P. (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. Cell *107*, 537–548.
13. Chen, X.B., Melchionna, R., Denis, C.M., Gaillard, P.H., Blasina, A., Van de Weyer, I., Boddy, M.N., Russell, P., Vialard, J., and McGowan, C.H. (2001). Human Mus81-associated endonuclease cleaves Holliday junctions in vitro. Mol. Cell *8*, 1117–1127.
14. Kaliraman, V., Mullen, J.R., Fricke, W.M., Bastin-Shanower, S.A., and Brill, S.J. (2001). Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. Genes Dev. *15*, 2730–2740.
15. Doe, C.L., Ahn, J.S., Dixon, J., and Whitby, M.C. (2002). Mus81-Eme1 and Rqh1 involvement in processing stalled and collapsed replication forks. J. Biol. Chem. *277*, 32753–32759.
16. Constantinou, A., Chen, X.B., McGowan, C.H., and West, S.C. (2002). Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. EMBO J. *21*, 5577–5585.
17. Aravind, L., Walker, D.R., and Koonin, E.V. (1999). Conserved domains in DNA repair proteins and evolution of repair systems. Nucleic Acids Res. *27*, 1223–1242.
18. Enzlin, J.H., and Scharer, O.D. (2002). The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. EMBO J. *21*, 2045–2053.
19. Park, C.H., Bessho, T., Matsunaga, T., and Sancar, A. (1995). Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. J. Biol. Chem. *270*, 22657–22660.
20. Bardwell, L., Cooper, A.J., and Friedberg, E.C. (1992). Stable and specific association between the yeast recombination and DNA repair proteins RAD1 and RAD10 in vitro. Mol. Cell. Biol. *12*, 3041–3049.
21. Biggerstaff, M., Szymkowski, D.E., and Wood, R.D. (1993). Correction of the ERCC1, ERCC4 and xeroderma pigmentosum group F DNA repair defects in vitro. EMBO J. *12*, 3685–3692.
22. Reardon, J.T., Thompson, L.H., and Sancar, A. (1993). Excision repair in man and the molecular basis of xeroderma pigmentosum syndrome. Cold Spring Harb. Symp. Quant. Biol. *58*, 605–617.
23. van Vuuren, A.J., Appeldoorn, E., Odijk, H., Yasui, A., Jaspers, N.G., Bootsma, D., and Hoeijmakers, J.H. (1993). Evidence for a repair enzyme complex involving ERCC1 and complementing activities of ERCC4, ERCC11 and xeroderma pigmentosum group F. EMBO J. *12*, 3693–3701.
24. Yagi, T., Wood, R.D., and Takebe, H. (1997). A low content of ERCC1 and a 120 kDa protein is a frequent feature of group F xeroderma pigmentosum fibroblast cells. Mutagenesis *12*, 41–44.
25. Komori, K., Fujikane, R., Shinagawa, H., and Ishino, Y. (2002). Novel endonuclease in Archaea cleaving DNA with various branched structure. Genes Genet. Syst. *77*, 227–241.
26. Sgouros, J., Gaillard, P.H., and Wood, R.D. (1999). A relationship between a DNA-repair/recombination nuclease family and archaeal helicases. Trends Biochem. Sci. *24*, 95–97.
27. Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.
28. Pingoud, A., and Jeltsch, A. (2001). Structure and function of type II restriction endonucleases. Nucleic Acids Res. *29*, 3705–3727.
29. de Laat, W.L., Sijbers, A.M., Odijk, H., Jaspers, N.G., and Hoeijmakers, J.H. (1998). Mapping of interaction domains between human repair proteins ERCC1 and XPF. Nucleic Acids Res. *26*, 4146–4152.
30. Tsutakawa, S.E., Muto, T., Kawate, T., Jingami, H., Kunishima, N., Ariyoshi, M., Kohda, D., Nakagawa, M., and Morikawa, K. (1999). Crystallographic and functional studies of very short patch repair endonuclease. Mol. Cell *3*, 621–628.
31. Bond, C.S., Kvaratskhelia, M., Richard, D., White, M.F., and Hunter, W.N. (2001). Structure of Hjc, a Holliday junction resolvase, from *Sulfolobus solfataricus*. Proc. Natl. Acad. Sci. USA *98*, 5509–5514.
32. Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y., and Morikawa, K. (2001). Crystal structure of the archaeal Holliday junction resolvase Hjc and implications for DNA recognition. Structure *9*, 197–204.
33. Daiyasu, H., Komori, K., Sakae, S., Ishino, Y., and Toh, H. (2000). Hjc resolvase is a distantly related member of the type II restriction endonuclease family. Nucleic Acids Res. *28*, 4540–4543.
34. Komori, K., Sakae, S., Daiyasu, H., Toh, H., Morikawa, K., Shinagawa, H., and Ishino, Y. (2000). Mutational analysis of the *Pyrococcus furiosus* Holliday junction resolvase hjc revealed functionally important residues for dimer formation, junction DNA binding, and cleavage activities. J. Biol. Chem. *275*, 40385–40391.
35. Kvaratskhelia, M., Wardleworth, B.N., Norman, D.G., and White, M.F. (2000). A conserved nuclease domain in the archaeal Holliday junction resolving enzyme Hjc. J. Biol. Chem. *275*, 25540–25546.
36. Gaillard, P.H., and Wood, R.D. (2001). Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. Nucleic Acids Res. *29*, 872–879.
37. Li, L., Elledge, S.J., Peterson, C.A., Bales, E.S., and Legerski, R.J. (1994). Specific association between the human DNA repair proteins XPA and ERCC1. Proc. Natl. Acad. Sci. USA *91*, 5012–5016.
38. Park, C.H., and Sancar, A. (1994). Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins. Proc. Natl. Acad. Sci. USA *91*, 5017–5021.
39. Sijbers, A.M., van Voorst Vader, P.C., Snoek, J.W., Raams, A., Jaspers, N.G., and Kleijer, W.J. (1998). Homozygous R788W point mutation in the XPF gene of a patient with xeroderma pigmentosum and late-onset neurologic disease. J. Invest. Dermatol. *110*, 832–836.
40. Otwinowski, Z., and Minor, W. (1997). Processing of X-Ray Diffraction Data Collected in Oscillation Mode, J. Charles, W. Carter, and R.M. Sweet, eds. (New York: Academic Press).
41. Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and

MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.
42. Fortelle, E.D.L., and Bricogne, G. (1997). Maximum likelihood heavy atom refinement for multiple isomorphous replacement and multianomalous diffraction methods. Methods Enzymol. *276*, 472–494.
43. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.
44. Nicholls, A. (1993). GRASP: Graphical Representation and Analysis of Surface Properties (New York: Columbia University).

Accession Numbers

The atomic coordinates and structure factors of the HefC547′ nuclease domain and its cocrystal structures have been deposited in the Protein Data Bank, as follows: 1j23, native; 1j24, Ca cocrystal; 1j24, Mn cocrystal; 1j22, Se-met.